

BEHAVIORAL VARIABILITY AND FREQUENCY-DEPENDENT SELECTION

ARMANDO MACHADO

DUKE UNIVERSITY

In Experiment 1, two conditions were compared: (a) a variability schedule in which food reinforcement was delivered for the fourth peck in a sequence that differed from the preceding N four-peck sequences, with the value of N continuously adjusted to maintain reinforcement probability approximately constant; and (b) a control condition in which the variability constraint was dropped but reinforcement probability remained constant. Pigeons responded approximately randomly under the variability schedule but showed strong stereotyped behavior under the control condition. Experiments 2 and 3 tested the idea that variability is the outcome of a type of frequency-dependent selection, namely differential reinforcement of infrequent behavior patterns. The results showed that pigeons alternate when frequency-dependent selection is applied to single pecks because alternation is an easy-to-learn stable pattern that satisfies the frequency-dependent condition. Nevertheless, 2 of 4 pigeons showed random behavior when frequency-dependent selection was applied to two pecks, even though double alternation is a permissible and stable stereotype under these conditions. It appears that random behavior results when pigeons are unable to acquire the stable stereotyped behavior under a given frequency-dependent schedule.

Key words: behavioral variability, frequency-dependent selection, Bernoulli process, Markov chain, lag analysis, response pattern, key peck, pigeon

Three lines of evidence show that response variability can be modulated by reinforcement contingencies. First, when pigeons, rats, and human subjects are given reinforcers for emitting variable behavior, variability increases (Blough, 1966; Machado, 1989; Morris, 1987; Neuringer, 1986; Page & Neuringer, 1985; Pryor, Haag, & O'Reilly, 1969; Schoenfeld, Harris, & Farmer, 1966; Van Hest, Van Haaren, & Van De Poll, 1989). Second, the more variability required to obtain a reinforcer, the more is typically obtained (Machado, 1989; Morris, 1989; Page & Neuringer, 1985). Third, response variability can be placed under stimulus control (Neuringer, 1991; Page & Neuringer, 1985). These three data sets

suggest that variability is another operant dimension of behavior, with the same theoretical status as response force, duration, and latency. The present study was designed to develop and test a theory that identifies (a) necessary and sufficient conditions to produce operant variability and (b) the processes or mechanisms through which these conditions generate variability.

What gives rise to variable performance? In a variability schedule, a sequence of, say, four responses distributed over two keys is reinforced provided an identical four-response sequence has not occurred during any of the last N trials (e.g., Page & Neuringer, 1985). On any given trial, then, the set of all possible sequences can be divided into two parts: the recent sequences, that is, the sequences that occurred during the last N trials, and the remaining, nonrecent sequences. Variability schedules continuously reinforce the nonrecent sequences. In addition, increasing N intensifies this form of nonrecent selection because, on average, the larger the value of N , the smaller the set of nonrecent sequences. On the other hand, and this is the critical observation, of two sequences with different probabilities of occurrence, the more probable sequence is more likely to be in the set of recent sequences, and, consequently, less likely to be reinforced. It

I thank John Staddon and the other L.A.B. group members for their continuous help and encouragement. The research was partially supported by grants to Duke University from NSF and NIMH (John Staddon, Principal Investigator). Experiment 1 was carried out while the author was at the Universidade de Lisboa, Portugal (my appreciation to the students who helped run the experiment). Parts of the paper were presented at the annual meeting of the Association for Behavior Analysis, Nashville, May 1990. Derivations of the functions represented in Figure 6 are available from the author upon request. Correspondence concerning this article may be sent to Armando Machado, Department of Psychology, Duke University, Durham, North Carolina 27706.

follows that variability schedules implement a particular type of *frequency-dependent selection*: They confer an increased advantage to the momentarily least probable, or least frequent, sequences of behavior.

The preceding analysis suggested the general hypothesis of this study: Variability is not directly or automatically reinforced; rather, it emerges from, or is mediated by, frequency-dependent selection. This hypothesis was implicit in a few previous studies. Blough (1966) used a frequency-dependent schedule with pigeons to obtain close-to-random interresponse times, and Shimp (1967) applied a similar procedure to patterns of key pecks.

The analysis assumes that sequences of four responses, used by the experimenter to define the variability schedule, are the only targets of frequency-dependent selection. Nonetheless, the following argument shows that patterns of other lengths are also targets of selection. Consider, for example, a pigeon that pecks the right key with probability p and the left key with probability $1 - p$. When $p > .5$, the majority of sequences emitted by this bird will contain more right than left pecks. It follows that those few sequences that contain more left than right pecks will be less recent on average and, consequently, will be reinforced more frequently. If the bird's response mechanism is sensitive to this gradient of reinforcement, then p should decrease. However, when $p < .5$, a symmetrical argument shows that p should increase. The frequency-dependent selection process will eventually stabilize when the probability of pecking right (or left) equals .5, in which case the amount of variation is maximal. In this example, reinforcement is balancing the strength of one-response patterns (individual left and right pecks), and sequence variability is an emergent property of the process.

The argument presented above for four- and one-response patterns is easily extended to two- and three-response patterns. As the results obtained with variability schedules attest, this multitarget selection property is clearly a *sufficient* condition to engender variable responding; is it also a *necessary* condition? Stated differently, how far in terms of pattern length does a frequency-dependent schedule need to go in order to generate variable behavior? This was the specific question addressed by the present set of experiments.

Experiment 2 (the reasons for, and the content of, Experiment 1 will be presented shortly) tested the simplest hypothesis, namely that sequence variability is obtained when individual responses are the only targets of frequency-dependent selection. The results did not support the hypothesis. Therefore, in Experiment 3, the frequency-dependent schedule was extended to two-response patterns.

As explained below, the dependent variables used in previous variability studies (e.g., Machado, 1989; Morris, 1987, 1989; Page & Neuringer, 1985) were not appropriate for Experiments 2 and 3. Therefore, the first experiment used a variability schedule with new measures of response variation. The results obtained in this experiment constituted a sort of baseline against which the data from Experiments 2 and 3 could be compared.

EXPERIMENT 1

The main purpose of the first experiment was to obtain descriptions of variable responding that could also be used in Experiments 2 and 3. Two common measures of response variability are the percentage of sequences meeting the variability requirement (i.e., sequences that differ from the last N sequences) and uncertainty measures derived from information theory (e.g., Machado, 1989; Page & Neuringer, 1985). The former cannot be used in experiments that do not specify a variability requirement (like Experiments 2 and 3 below) or in experiments that use a nonconstant N criterion (e.g., Machado, 1989). More generally, measures based on sequences as units of analysis are not useful in experiments that analyze behavior in units other than the sequence. On the other hand, uncertainty, equal to $-\sum p_i \log_2(p_i)$ where p_i is the probability of sequence i , is also inadequate because it is a composite measure that does not reveal which particular type of deviation from random performance has occurred (e.g., the same uncertainty score can represent either a tendency to avoid or a tendency to repeat the last response). To compare variability schedules with other frequency-dependent schedules, it is necessary to use more descriptive, less procedure-dependent measures of the subject's actual behavior.

Four different measures of variability were used in the present experiment: (a) the percentage of pecks on the right key, (b) the per-

centage of response pairs involving alternations (left-right and right-left pecks), (c) a set of conditional probabilities known as lag analysis, and (d) statistical tests based on the theory of Markov chains.

The schedule used in the first experiment differed from the variability schedules used by other investigators, in that the variability requirement, N , was not fixed but changed continuously in order to keep constant the probability of reinforcement per trial. A problem with a fixed variability requirement is that the experimenter loses control over the probability, hence the frequency and intermittency, of reinforcement; as the degree of response variability changes, the obtained frequency of reinforcement also changes. The effects on behavioral variability of the operant contingency will then be confounded with the well-known effects of intermittent but nondifferential reinforcement (see Boulanger, Ingebos, Lahak, Machado, & Richelle, 1987, for a review of the effects of intermittent reinforcement on response variability).

The percentile schedule adapted by Machado (1989) from Platt (1973) was used because it allows the differential reinforcement of variable patterns while keeping constant the probability of reinforcement per trial. In an ABA design, reinforcement was contingent on sequence variation during the A phases, whereas in Phase B variability was permitted but not required. Throughout, the frequency of reinforcement was kept constant.

METHOD

Subjects

Five experimentally naive homing pigeons (*Columba livia*) were used. Each pigeon was maintained at 80% (± 15 g) of its free-feeding body weight. Water and grit were continuously available in the home cage.

Apparatus

The experimental chamber was 32 cm along the sides and 45 cm high. The floor was wire mesh and all the walls (except the one on which the keys and feeder were mounted) and the ceiling were Plexiglas. The box was placed 50 cm above the floor in a small cubicle 1.5 m wide and 2.5 m high. A 40-W white house-light, permanently lit, was located on the ceiling of the cubicle. The front wall was equipped

with two response keys (2.2 cm diameter) symmetrically located 2.5 cm to either side of the midline 23 cm from the floor. A force of 0.2 N on either response key operated a micro-switch. Each key could be illuminated with a 5-W orange light. Directly below the keys, a hopper opening (4.5 cm by 7 cm and 7 cm from the floor) permitted access to mixed grain. A 7.5-W white light illuminated the hopper when grain was delivered. All events were controlled and data were recorded by a Commodore® 64 computer. Data were later transferred to a larger computer for analyses.

Procedure

Pretraining. Sessions were conducted 6 days per week at approximately the same time each day. All pigeons were trained to peck the keys under a modified autoshaping procedure developed by Schwartz (1980). After variable intervals averaging 60 s, one or both keys were randomly selected and lit for 6 s, after which reinforcement (4-s access to grain) was delivered. During food presentations, both key-lights were turned off. If a peck occurred while a key was illuminated, food was presented immediately. Sessions ended after 50 reinforcers were delivered.

The next four sessions constituted a transition phase used to adapt the birds to intermittent reinforcement. Each session was divided into trials, and each trial was as follows: At the beginning of the trial, one randomly selected key was lit. A peck on the lit key turned off the keylight for a 1-s interpeck interval, during which any key peck reset the timer for the interval. Following the interpeck interval, one key was again lit randomly. After four pecks had been emitted, the trial ended and either a 4-s reinforcer or a 1-s intertrial interval (ITI) followed immediately (note that the fourth peck was not followed by the interpeck interval). During both the ITI and food presentations the keylights were off, but the houselight remained illuminated. As in the interpeck interval, a peck during the ITI reset the timer. In summary, a trial involved the following sequence of events: Each of the first three pecks was followed by an interpeck interval, and the fourth peck was followed either by food or by an ITI. Immediately after the ITI or the food presentation, another trial began. Across the four sessions of the transition phase, the probability of reinforcement per trial

decreased from 1 to .4, and the number of trials increased from 50 to 100.

After the transition phase the experiment proper began. All procedural details remained the same as in the transition phase, except that (a) at the beginning of each trial both keylights were illuminated simultaneously and (b) the reinforcement probability was controlled by a percentile variability schedule (see below).

The percentile variability schedule. In order to apply a percentile schedule (Platt, 1973; see also Machado, 1989) to sequence variability, some indication of the subject's current tendency to vary its behavior is needed. This is achieved as follows: Once a sequence of four pecks is produced, it gets a *recurrence time* equal to the number of sequences intervening between its last occurrence and its present recurrence. For example, suppose a pigeon produced on successive trials the following sequences of left (L) and right (R) pecks: LLRL, RLL, RRRR, LLRL, and RRRR. The last sequence, RRRR, would get a recurrence time equal to 1, and the preceding sequence, LLRL, would get a recurrence time equal to 2. Thus, each sequence gets a recurrence time that measures its recency, and the frequency distribution of the last x recurrence times (where x is a schedule parameter) provides an indication of the subject's current tendency to vary its responses (e.g., if most recurrence times are small this means that the subject is emitting only a small subset of all possible sequences and therefore its behavior is not highly variable).

From the last x recurrence times, a percentile is computed and its value becomes the criterion N for the next trial. For example, if $x = 20$ and the 70th percentile is being used, the criterion N is found by first ordering from low to high the last 20 recurrence times and then determining the score at the 14th ($20 \times .7$) rank. If the recurrence time of the sequence emitted on the next trial is greater than N , the sequence is eligible for reinforcement. Finally, the recurrence time of the new sequence replaces the one that occurred x trials before; as in a shaping procedure, the criterion depends only on the most recent behavior of the subject.

One aspect of the percentile schedule is critical: Although the cardinal value of the criterion, N , may change from trial to trial, the probability of emitting a sequence whose recurrence time is greater than the criterion (i.e.,

a reinforceable sequence) is always constant and equal to the complement of the percentile. This happens because sample percentiles estimate population percentiles and, as rank statistics, percentiles do not depend on the particular distribution of sequences or recurrence times. Thus, in the example above with the 70th percentile, the probability of a reinforceable sequence would always be .3.

A potential problem of the percentile variability schedule is that, occasionally, the variability requirement cannot be met. To illustrate this situation, suppose that $N = 30$ in the current trial. This means that the next sequence should differ from the last 30 sequences in order to be reinforceable. Nonetheless, if all possible sequences were produced on the last 30 trials, then no sequence can get a recurrence time greater than 30. Whenever the computer detects this situation, the criterion is reduced to the next value that makes possible the occurrence of a reinforceable sequence. In the example, N could be reset to 20 because one sequence could then be produced and reinforced. In the present experiment, this situation was very rare and did not affect the overall workings of the schedule.

In both A conditions of an ABA design, each sequence of four pecks was given a recurrence time, with a maximum value of 50 (recurrence times have no upper bound). If this recurrence time was greater than the current criterion N , reinforcement was delivered; otherwise, the ITI followed. The criterion for the next trial matched the 70th percentile of the last 20 recurrence times. It follows that, on each trial, the probability of exceeding the criterion was .3. Because all (and only those) sequences meeting the criterion were reinforced, the probability of reinforcement per trial was also .3.

During Condition B, no criterion was used and every sequence was eligible for reinforcement (equivalently, one might say that the criterion matched the "zero percentile"). To have the same probability of reinforcement per trial as in Condition A, each sequence was reinforced with a probability of .3. In summary, in Condition A (the variability condition), reinforcers followed variable sequences, whereas in Condition B (the no-variability condition), reinforcers were not contingent on sequence variation. The probability of reinforcement per trial always equaled .3.

At the beginning of each session, the last trials of the previous session were used to compute the criterion and determine the initial recurrence times. Sessions ended after 100 trials. Each condition was in effect until the percentages of pecks on the right key and alternations showed no clear trend for at least five consecutive sessions. In the last phase, practical difficulties prevented Subjects 3 and 5 from participating in additional sessions and, consequently, their performance was not stable by the end of training.

Data Analysis

Although the reinforcement rule was defined in terms of sequences of four responses, the data analysis was not based on sequences but on the string of 400 individual pecks (100 trials times four pecks per trial) emitted by each subject in each session. In addition, whenever pairs, triplets, or, in general, n -tuplets of responses are mentioned, it should be understood that a moving window was used to define those patterns. Hence, in the string RRLRL, the first two pairs are RR and RL, and the first two triplets are RRL and RLR.

The percentage of pecks on the right key and the percentage of response pairs that involved alternations (i.e., RL + LR) were used as molar measures of response variability. More molecular measures were provided by lag analysis and Markov chains. Lag analysis determines how the current response affects the probability of subsequent responses of the same type. It proceeds as follows: For each session, the probability of pecking, say, the right key is estimated from the relative frequency of right pecks. This *unconditional* probability is defined as Lag 0. The *conditional* probability of a right peck given that the previous response was also a peck on the right key is defined as Lag 1. Generalizing, lag n is the probability of pecking right given a right peck n responses before and irrespective of the (left or right) location of the intervening $n - 1$ responses. If the data contain no sequential dependencies, that is, if the n th response is independent of the preceding $n - 1$, $n - 2$, . . . , responses (for all n), then the conditional probabilities (Lags 1, 2, . . .) will all be close to the unconditional probability (Lag 0). On the other hand, if a bird avoids the last response and pecks frequently in an alternating pattern (RLRLR . . .), then Lag 0 will be close to .5, odd lags (1, 3, 5, . . .) will

be systematically lower than Lag 0, and even lags (2, 4, 6, . . .) will be systematically greater than Lag 0. A zigzag profile of conditional probabilities will be generated. Other profiles are obviously possible and, in general, a close inspection of these profiles is all that is required to determine the particular structure of the data. Bakeman and Gottman (1986) and Gottman and Roy (1990) may be consulted for further details of lag analysis.

Lag analysis estimates the effect of a single response on the probability of subsequent responses. However, the probability of the next response may also depend on all of the last n responses. This possibility can be tested by measuring how well a Markov chain model of order n fits the data. A Markov model of order n assumes that the probability of the next response depends exclusively on the previous n responses. When testing a model of order n , the null hypothesis is a model of order $n - 1$, and the test simply compares the predictions based on the null hypothesis against the data. The following example describes the logic of the test and its implementation. Suppose the investigator wants to fit a model of order $n = 2$; the null hypothesis is then a model of Order 1 ($n - 1$); that is, the current response depends stochastically only on the previous one. Then, under the null hypothesis, the probability of each response triplet (sequence of length $n + 1$) can be computed as follows (the example generalizes to all triplets):

$$\begin{aligned} P(\text{RRL}) &= P(\text{RR})P(\text{L}|\text{RR}) \\ &= P(\text{RR})P(\text{L}|\text{R}), \end{aligned}$$

where the second equality follows because the assumption is that only the preceding response affects the probability of the current response. Both probabilities, $P(\text{RR})$ and $P(\text{L}|\text{R})$, are estimated from the data, and a chi-square test with $2^{n-1} = 2$ degrees of freedom determines whether the discrepancy between predicted and observed frequencies of triplets is large enough to reject the null hypothesis. Note that information about pairs is used to predict frequencies of triplets. If the null hypothesis is not rejected, then the information contained in the pairs is sufficient to predict the triplets. Therefore the investigator asks next if the information contained in single responses [i.e., $P(\text{R})$ and $P(\text{L})$] is also sufficient to predict the frequency of each pair. In other words, the test is repeated but this time $n = 1$. Had the null

hypothesis been rejected, then a new test with $n = 3$ determines whether the fitting of the model improves as one goes from Order 2 to Order 3. In summary, the tests are applied successively (either by increasing or decreasing n) until the order of the chain is determined. This order is an index of the extent of sequential dependencies present in the data. Gottman and Roy (1990) give additional information on Markov chains and associated tests.

RESULTS AND DISCUSSION

Figure 1 shows the percentage of pecks on the right key, the percentage of alternations, and the percentage of reinforced pecks for each subject and session of the experiment. For all birds, the percentage of pecks on the right key was close to 50% at the end of both variability phases but approached 0% or 100% during the no-variability phase. Alternation increased to values close to 50% when variability was required but decreased steadily when reinforcers were delivered without regard to response variability. These differences in response variability across conditions were not due to differences in the frequency of reinforcement. As Figure 1 shows, the percentage of reinforced pecks per session was approximately constant and close to the scheduled value.

When a bird strongly prefers one response key (as in the no-variability condition), the analysis of higher order patterns is somewhat superfluous because the results of such an analysis are well predicted by the probabilities of each individual response. In the extreme case of exclusive preference (e.g., Birds 1 and 2, no-variability condition), the concept of higher order patterns itself becomes meaningless. On the other hand, when individual responses occur equally (as in the variability condition), further analyses are needed to ascertain whether any higher order stereotypies developed. Thus, lag analysis and Markov tests were applied solely to the data from the last five sessions of the two variability conditions.

Figure 2 shows the results of the lag analysis (up to Lag 5) for right pecks (similar results were obtained for left pecks). The first point of each profile and the horizontal line segment represent the unconditional probability of pecking the right key. The next point, Lag 1, represents the probability of a right peck given that the previous response was also a right

peck. The subsequent points have similar interpretations; for example, the last point, Lag 5, represents the conditional probability that the fifth response after a right peck was also a right peck. Strong deviations from the horizontal line segment indicate sequential dependencies.

In general, the conditional probabilities were close to the unconditional probability (i.e., Lag 0); this result suggests that successive key pecks were stochastically independent. In a few sessions, some subjects tended either to alternate or to repeat the last peck. The data from Bird 3 illustrate both types of deviation from random performance: The arrow pointing left shows the switching pattern, with Lag 1 much lower than Lag 0 and Lag 2 much greater than Lag 0; the right arrow shows the repeating pattern, with Lag 1 much greater than Lag 0. However, these profiles were not consistently displayed across sessions. Bird 5 was the exception, in that in most of its sessions the profiles corresponding to the strategies RLRL... (right arrow) or RLLRLL... (left arrow) were clearly shown.

Chi-square tests based on the Markovian hypothesis were performed for each of the last five sessions of both variability conditions. Table 1 displays the results for first- and second-order models. A significant difference for a first-order model indicates that the location of the next peck depended stochastically on the location of the last peck. Similarly, a significant difference for a second-order model indicates that the next peck depended on the last two pecks. The results of the tests agreed closely with the lag data: For Birds 1, 2, 3, and 4, only seven of 40 tests provided significant evidence for a first-order dependency and only one provided evidence for a second-order dependency (throughout, $\alpha = .01$). No test suggested a third-order dependency. Hence, in most sessions, no sequential dependencies between responses were apparent. In those few sessions in which sequential dependencies did develop, they were mainly first order; that is, the next response depended stochastically only on the previous response. As in the lag analysis, Subject 5 was the exception: In the first variability condition, either a first- or a second-order dependency was always present. Although a first-order dependency was also evident in three of the last five sessions of the last condition, these results must take into account the

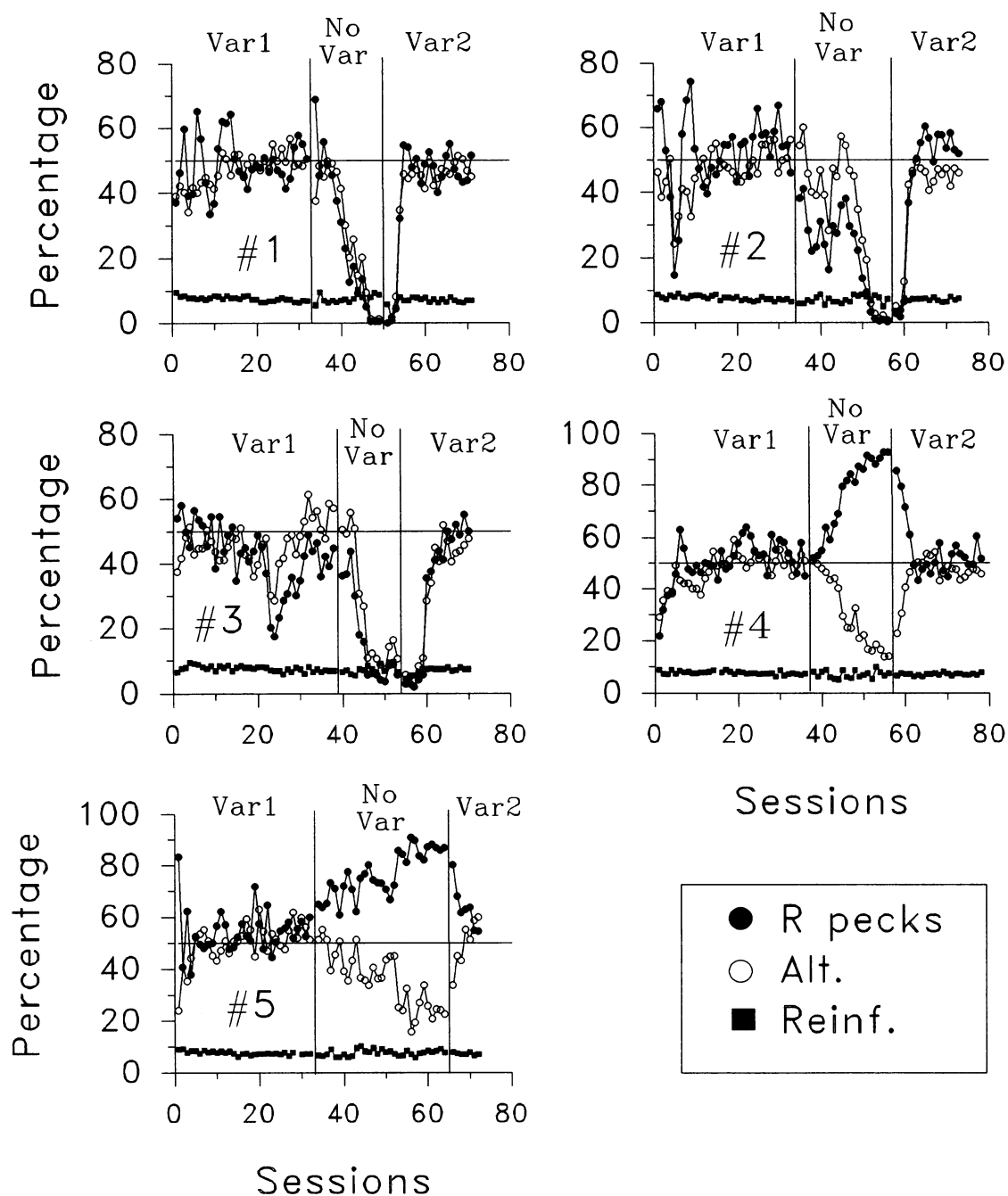


Fig. 1. Percentages of pecks on the right key, alternations (LR + RL pairs), and reinforced pecks for each bird in each session. In both "var" conditions, reinforcement was contingent on sequence variability. In the "no-var" condition, variability was permitted but not required for reinforcement.

fact that this bird's performance had not reached stability when the experiment ended.

The present findings show that contingent reinforcement can engender and maintain re-

sponse variability, and they agree with previous findings reported in the literature (Machado, 1989; Morris, 1987, 1989; Page & Neuringer, 1985; Van Hest et al., 1989). The

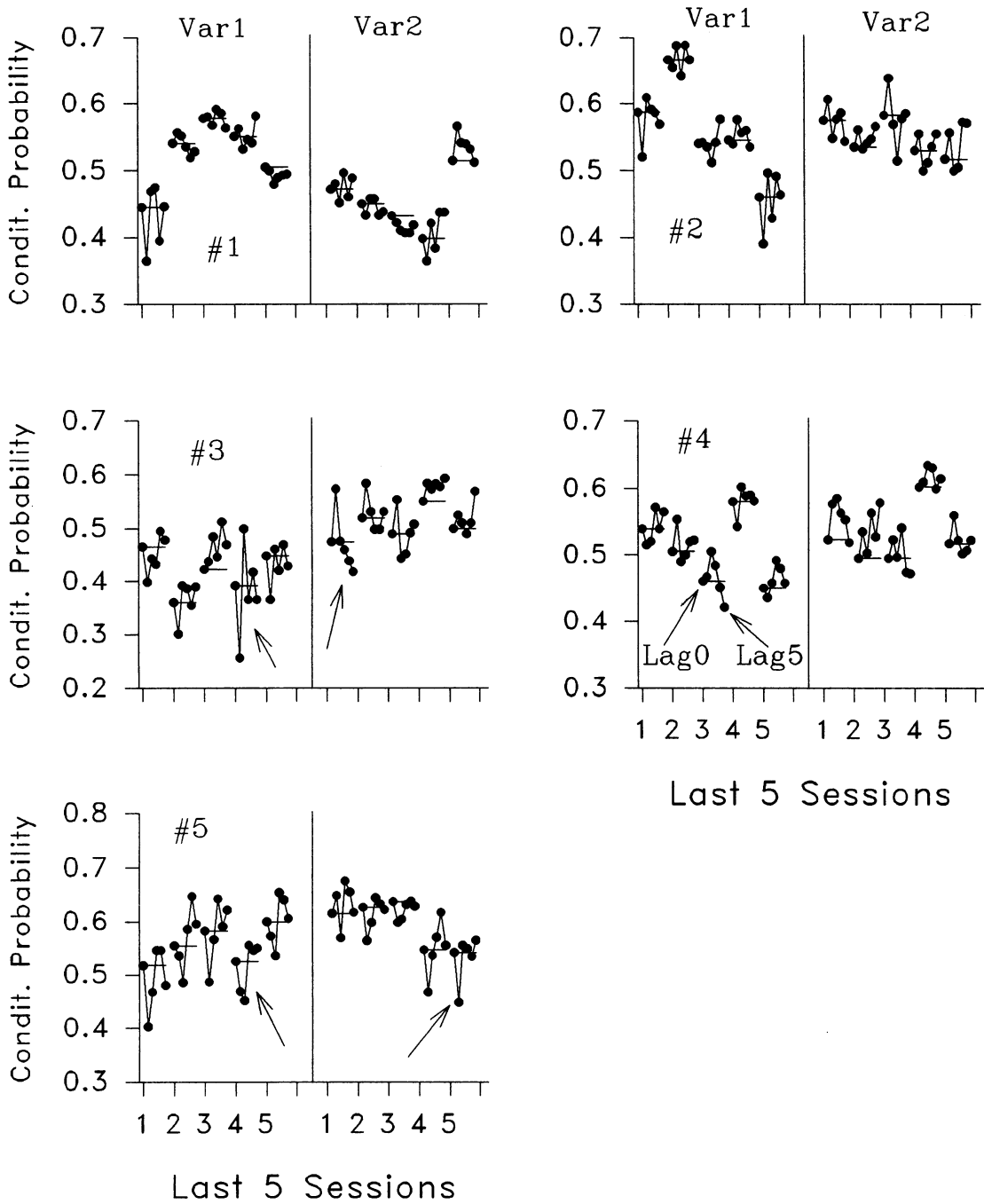


Fig. 2. Conditional probability profiles for the last five sessions of the two variability conditions ("Var 1" and "Var 2"). Each set of six connected points, Lags 0 to 5, corresponds to one session. Strong deviations from the horizontal line segment, equal to the unconditional probability $P(R)$, indicate sequential dependencies. The probability profiles marked with arrows are discussed in the text.

Table 1

Chi-square values for a first- and second-order Markov chain model in the last five sessions of each variability condition. For a significance level of .01, critical values are $\chi^2(1) = 6.63$ and $\chi^2(2) = 9.21$ for tests of a first- and a second-order model, respectively.

Condi- tion	Ses- sion	Bird 1		Bird 2		Bird 3		Bird 4		Bird 5	
		1st	2nd	1st	2nd	1st	2nd	1st	2nd	1st	2nd
Var 1	1	8.27*	2.24	10.12*	1.33	6.31	7.09	0.87	3.69	22.41*	11.40*
	2	0.31	2.22	0.39	1.79	3.33	4.19	4.02	1.12	0.96	11.66*
	3	0.00	0.56	0.01	0.89	0.21	4.93	0.04	2.88	20.64*	4.32
	4	0.28	0.81	0.07	1.83	19.78*	6.65	3.18	11.04*	6.04	13.22*
	5	0.05	0.05	6.48	1.22	8.77*	1.03	0.26	1.98	1.54	13.70*
Var 2	1	0.13	0.13	1.96	2.10	14.97*	1.82	4.73	8.29	3.16	6.41
	2	0.37	0.37	0.99	0.10	7.76*	1.26	2.58	0.71	12.34*	5.12
	3	0.20	0.20	8.05*	1.00	5.76	6.15	1.21	0.50	4.52	3.95
	4	1.33	1.34	1.09	4.24	2.36	1.45	0.05	2.19	12.41*	2.47
	5	4.36	4.37	2.83	4.23	0.81	1.59	2.80	2.39	16.69*	1.24

* $p < .01$.

present experiment also replicates, with a within-subject design, Machado's (1989) experiment that used a between-subjects design. Because the results were identical in both cases, the present experiment increases the generality of the earlier findings concerning the effects of operant contingencies on response variability when the frequency of reinforcement is kept constant.

Taken together, present and past results suggest that under variability schedules the mechanism that generates key pecks becomes formally equivalent to a Bernoulli process. That is, subjects behave as if they were flipping a coin before each response and acting accordingly. Successive responses are independently generated. Although the parameter of the Bernoulli process (the bias of the coin) fluctuates across sessions, it is generally close to .5. The few deviations from this description are mostly first order: a tendency to repeat or to avoid the last response. These behavioral effects define what the frequency-dependent selection, or any other hypothesis, must ultimately explain.

EXPERIMENT 2

The second experiment asked whether differential reinforcement of the nonpreferred response alternative is a sufficient condition to engender high levels of behavioral variability. In order to test this possibility, a frequency-dependent schedule that targeted only individual responses was designed as follows. Each individual peck constituted a trial (rather than four pecks as in Experiment 1), and from the

last m trials the momentary probability of pecking the right key was estimated. As this probability increased, the payoff probability for a right peck decreased and that for a left peck increased. Conversely, when right pecks were rare, the payoff probability for right pecks increased and that for left pecks decreased. The reinforcement probabilities for right and left pecks were equal only at indifference [i.e., $P(R) = P(L) = .5$], in which case, assuming response independence, the degree of sequence variability is at its maximum.

On each trial, the contingencies of reinforcement depended on the momentary probability of pecking the right key; this probability was estimated by the proportion of right pecks on the last m trials. How large should m be? If the momentary probability of pecking the right key is changing rapidly, then a small sample is preferable because it is more sensitive to this change. On the other hand, if the probability of a right peck is not changing, a large sample has the advantage of yielding more reliable estimates of that probability. Because the rate of change of the probability of pecking the right key was not known, different sample sizes were used across the phases of the experiment.

Another reason to vary the sample-size parameter comes from its similarity with the criterion N in variability schedules. A smaller criterion in variability schedules implies that fewer (four-response) trials without the occurrence of a target sequence have to elapse before that sequence is again eligible for reinforcement. Similarly, in frequency-dependent schedules, a smaller sample m implies

that the probability of reinforcement for a given response increases more rapidly with successive (one-response) trials without that response. To illustrate, suppose that the last m pecks were on the right key. Then, the estimated probability of a right peck, $P(R)$, equals 1 (m right pecks out of m trials). If a left peck occurs on the next trial, $P(R)$ changes from 1 to $(m - 1)/m$. The difference between the two estimates equals $1/m$, which implies that after each left peck, $P(R)$ decreases at a faster rate when the sample size m is small than when it is large. Because the probability of reinforcement is inversely related to $P(R)$, it follows that the probability of reinforcement for a right peck increases more rapidly with each left peck when the sample is small. Given the similarity between the sample size m and the criterion N , previous findings (e.g., Machado, 1989; Page & Neuringer, 1985) suggested that smaller samples would also be associated with less response variability.

The frequency-dependent schedule used in Experiment 2 shared three additional properties with the previous percentile schedule: (a) Although the trial definition changed (four responses in Experiment 1, only one response in Experiment 2), the expected probability of reinforcement per trial was constant; (b) in Experiment 1 when a sequence was repeated continuously its probability of reinforcement decreased to zero. Similarly, in the present schedule, when an exclusive preference developed for one alternative the payoff probability for that alternative was zero, and (c) the contingencies of reinforcement did not favor one response over the other.

METHOD

Subjects

Four pigeons with previous experience with time-related reinforcement schedules were used. None of the birds had been exposed to frequency-dependent or variability schedules. The subjects were maintained at 80% (± 15 g) of their free-feeding body weights. Water and grit were continuously available in their home cages.

Apparatus

An outer box, with a houselight permanently lit on the ceiling and a ventilating fan on the front wall, enclosed the experimental

chamber. All other details were as in Experiment 1.

Procedure

Pretraining. For all birds, only one auto-shaping session was needed to reinstate key pecking. The transition phase, implemented for four sessions, was similar to that used in Experiment 1, with the difference that each trial required only one peck. Thus, at the beginning of each trial, one randomly selected key was lit. A peck on the lit key was followed either by reinforcement (4-s access to grain) or by the 1-s ITI. Any peck during the ITI reset the timer for the interval. Note that because only one peck was needed to end the trial, there was no interpeck interval. After the ITI or food presentation, the next trial began immediately. Across the four transition sessions, the probability of reinforcement after each peck was reduced from .15 to .05 and the number of trials increased from 50 to 440.

During the experiment proper all procedural details were as in the transition phase except that (a) both keys were lit at the beginning of each trial and (b) the reinforcement rule was frequency dependent (see next section).

The reinforcement schedule. After each trial, the proportion, s , of right pecks emitted on the last m trials was computed. This proportion estimated the momentary probability of pecking the right key. On the next trial, the probabilities of reinforcement for a right and a left peck, $P(\text{Ref}|R)$ and $P(\text{Ref}|L)$, respectively, were given by the two hyperbolic functions

$$P(\text{Ref}|R) = u(s) = \begin{cases} \frac{k(1-s)}{s}, & s \geq \frac{k}{1+k} \\ 1, & \text{otherwise} \end{cases} \quad (1)$$

$$P(\text{Ref}|L) = v(s) = \begin{cases} \frac{ks}{1-s}, & s \leq \frac{1}{1+k} \\ 1, & \text{otherwise} \end{cases}$$

plotted in Figure 3. The constant k represents the expected probability of reinforcement per trial, $P(\text{Ref})$, where

$$P(\text{Ref}) = P(\text{Ref}|R) \times P(R) + P(\text{Ref}|L) \times P(L).$$

The reinforcement rule has the following features: The payoff probability for one response varies inversely with its probability of

occurrence; in case of exclusive preference, the payoff probability for the preferred response equals zero; the expected probability of reinforcement per trial equals k whenever the proportion of right pecks is in the interval $[k/(k+1), 1/(k+1)]$; and $P(\text{Ref}|\text{R})$ and $P(\text{Ref}|\text{L})$ are symmetric around the line $s = .5$, which means that preferences for the left or the right key are handled similarly.

All birds were initially exposed to a simplified version of the reinforcement schedule defined by Equation 1. In this version, one of the functions, $P(\text{Ref}|\text{R})$ or $P(\text{Ref}|\text{L})$, was initially set to zero while the other remained unchanged. After a few sessions, the contingencies were reversed; that is, the function initially set to zero was then set to its value in Equation 1 and the other function was set to zero. After this initial training, the experiment proper began.

Across the three successive phases of the experiment, the sample size m equaled 40, 10, and 80, respectively. Each phase lasted for at least 15 sessions and until the relative frequencies of right pecks and alternations were judged stable using the same criterion as in Experiment 1. The expected probability of reinforcement per trial, k in Equation 1, was always .05. The first 40 pecks of each session were not eligible for reinforcement; they simply filled the memory window from which the probability of pecking the right key was calculated. However, in the last phase ($m = 80$) these initial 40 pecks were not enough to fill the window. Therefore, from Trials 41 up to 80, the window increased with each successive response. Sessions ended after 440 trials.

RESULTS AND DISCUSSION

The string of 440 individual pecks emitted by each bird in each session was analyzed as in Experiment 1. Figure 4 shows the percentage of pecks on the right key, the percentage of RL + LR pairs, and the number of reinforcements for each bird in each session. For all birds, the percentage of pecks on the right was very stable both within and between conditions, and its value was close to 50%.

Over the last five sessions when the sample size was 40, the average percentage of alternations ranged across birds from 42% to 60%. Changing the sample to 10 increased this average for all birds (range, 56% to 64%). Finally, when the sample size was 80, Subject

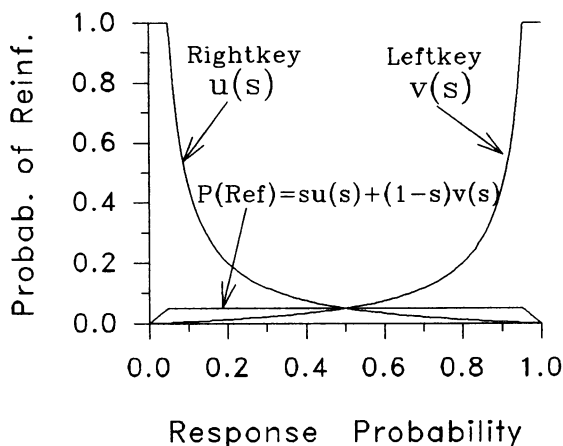


Fig. 3. Reinforcement probability for right and left pecks, as a function of the current probability of pecking the right key, s . The expected probability of reinforcement per trial, $P(\text{Ref})$, is constant in the interval $k/(k+1) \leq s \leq 1/(k+1)$ and equal to $k = .05$. See Equation 1 in text.

156 slightly decreased, whereas the remaining birds maintained, the frequency of alternations (range, 50% to 66%). Although the smallest sample yielded the highest degree of alternations for 3 subjects (except 170), the largest sample did not yield the lowest value.

A lag analysis was performed for each of the last five sessions of each condition. Figure 5 displays the results for all birds. The lag profiles show that, in general, the pigeons developed strong sequential dependencies between consecutive key pecks. When the sample size was 40, Birds 157 and 170 developed a reliable switching pattern (Lag 1 much lower than Lag 0), whereas Bird 156 showed a strong repeating pattern (Lags 1, 2, . . . much greater than Lag 0). Bird 151, however, did not display sharp sequential dependencies. When the sample size was 10, all birds developed a pattern of switching, consistently shown from session to session and, in some cases, remarkably strong. Not only was Lag 1 much lower than Lag 0, but Lag 2 was much greater than Lag 0 (except for Bird 156). The corresponding zigzag profiles shown in Figure 5 indicate that the alternation pattern became stronger during this phase of the experiment. Finally, when the sample size was 80, Bird 156 showed no strong sequential dependencies, whereas the other subjects maintained the switching pattern.

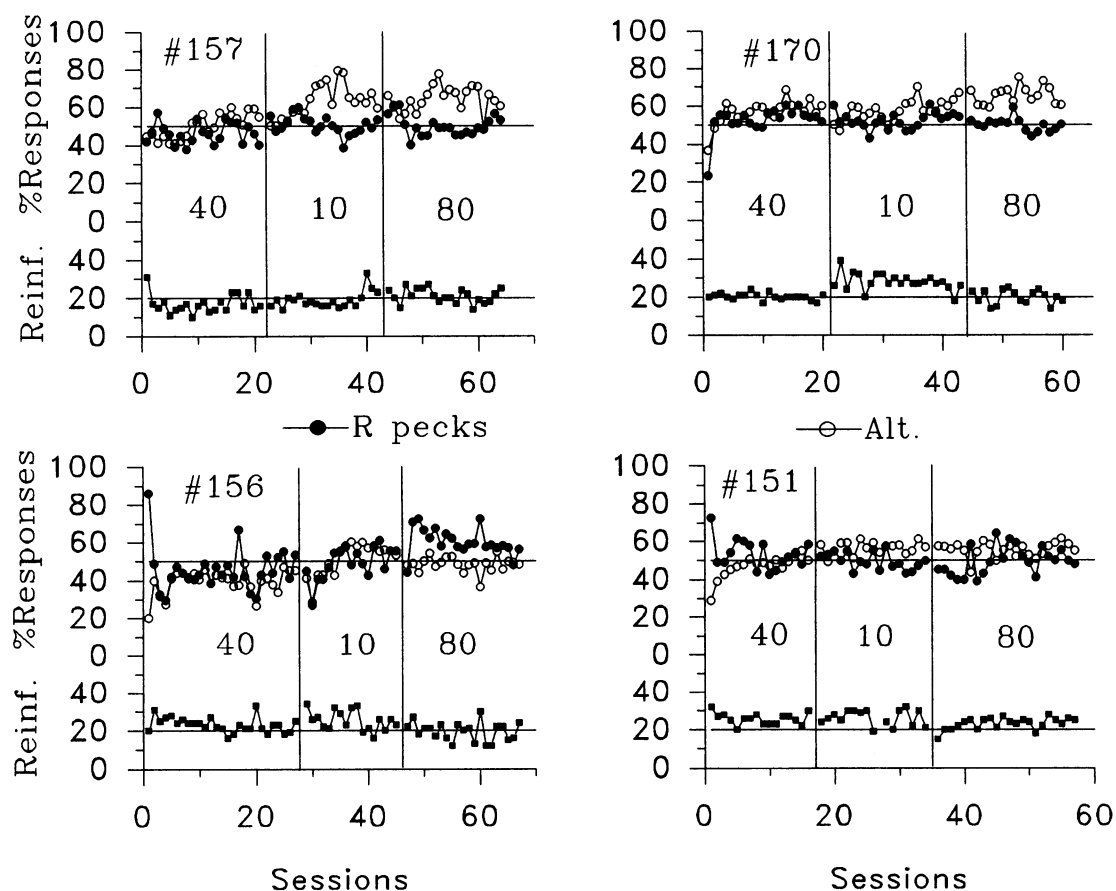


Fig. 4. Percentages of pecks on the right key and alternations (i.e., LR + RL), together with the number of reinforcers obtained in each session. The numbers in each condition show the sample size m from which the probability of pecking the right key was estimated.

Chi-square tests, assessing the goodness of fit of the Markov chain model, were also performed for each of the last five sessions of each condition. The results, shown in Table 2, indicate that a first-order dependency was present in 40 of the 60 sessions, but a second-order dependency was present in only 12 of the 60 sessions. No test supported a third-order dependency. Thus, for most sessions, the next response depended stochastically on only the preceding response.

These findings show that, irrespective of the sample size, the frequency-dependent schedule engendered the first feature of performance obtained with variability schedules, namely the identical frequency of right and left pecks. This similarity of results, however, hid two different response patterns: In the frequency-dependent schedule the birds alternated reliably, es-

pecially when the sample was small, whereas in variability schedules the pigeons behaved as if they were flipping a coin before pecking the keys. To be sure, switching patterns also occurred under variability schedules (e.g., Figure 2, Bird 3, left arrow), but not with the same intensity within sessions or reliability across sessions.

Why did the birds develop strong switching patterns in Experiment 2? For all birds, the proportion of reinforced alternations (Alt) was always greater than the proportion of reinforced repetitions (Rep). Averaging over all sessions and conditions showed that $P(\text{Ref}|\text{Alt})$ varied across birds from .05 to .08, whereas $P(\text{Ref}|\text{Rep})$ varied from .03 to .05. Hence, the schedule continuously favored alternations over repetitions. To understand this schedule effect (not obvious from Equation 1), expected prob-

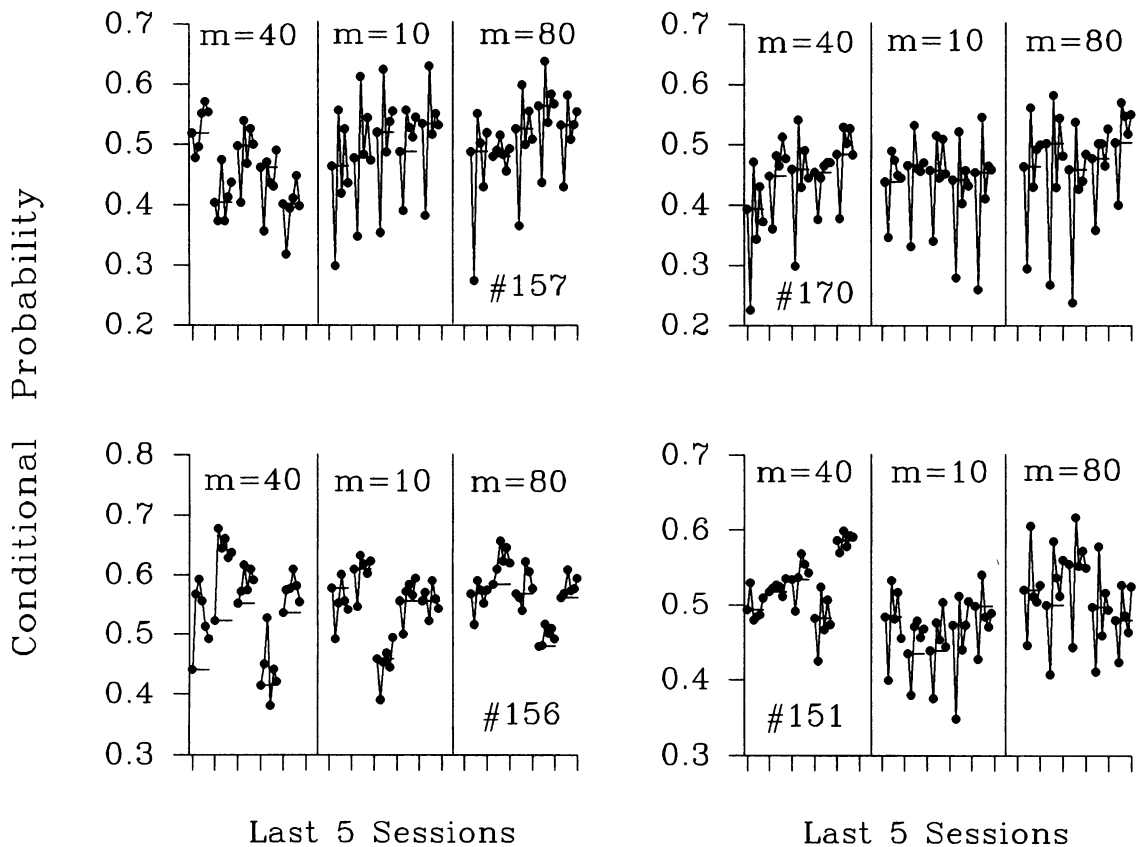


Fig. 5. Conditional probability profiles for the last five sessions of each condition. Values of m refer to the size of the sample from which the unconditional probability of pecking the right key was estimated. Other details as in Figure 2.

abilities of reinforcement were derived by assuming, as the results suggested, that the response mechanism was a first-order Markov chain. The last response, a left or a right peck, defined the current state of the chain, and the current state determined the response probabilities for the next trial. These probabilities, $P(R|R)$ and $P(L|L)$, and their complements, $P(L|R) = 1 - P(R|R)$ and $P(R|L) = 1 - P(L|L)$, fully characterize the behavior of the chain; throughout, they were assumed to be stationary. For the case $m = 10$, Figure 6 shows three resulting functions: the overall probability of reinforcement, $P(\text{Ref})$; the probability of reinforcement given an alternation, $P(\text{Ref}|\text{Alt})$; and the probability of reinforcement given a repetition, $P(\text{Ref}|\text{Rep})$. Qualitatively similar results were obtained for larger sample sizes.

The upper graph in Figure 6 shows that, when sample size is taken into account, the

overall probability of reinforcement will generally be greater than the scheduled value of $k = .05$ (except for extreme repeat probabilities). This result explains the finding that, in some conditions (e.g., Figure 4, Bird 170, $m = 10$), the obtained frequency of reinforcement was well above the predicted number of 20 ($.05 \times 400$ reinforceable trials). The graph also shows that the highest probability of reinforcement is achieved with high probabilities of repetitions, not switches, in clear contrast to the birds' performances.

The middle graph shows that $P(\text{Ref}|\text{Alt})$ is always greater than .05, and its range of variation, for different pairs of $P(R|R)$ and $P(L|L)$, is considerable; notice the different scale for the vertical axis. In contrast, the bottom graph shows that $P(\text{Ref}|\text{Rep})$ is usually below .05. It follows that, in the present schedule, whenever the subjects tended to repeat the last response, alternations were selectively re-

Table 2

Chi-square values for a first- and second-order Markov chain model in the last five sessions of each condition. For a significance level of .01, critical values are $\chi^2(1) = 6.63$ and $\chi^2(2) = 9.21$, for tests of a first- and a second-order model, respectively. m = sample size.

Condi- tion	Session	Bird 157		Bird 156		Bird 170		Bird 151	
		1st	2nd	1st	2nd	1st	2nd	1st	2nd
$m = 40$	1	3.44	1.92	22.55*	23.96*	32.70*	2.65	2.34	0.86
	2	1.20	6.08	45.71*	12.17*	11.55*	6.41	0.03	0.38
	3	16.13*	0.87	0.85	10.65*	38.75*	7.33	3.64	0.69
	4	16.55*	1.40	1.79	16.81*	9.49*	6.64	5.39	3.18
	5	9.32*	2.11	2.98	3.50	18.72*	3.42	0.56	0.56
$m = 10$	1	41.56*	4.07	17.14*	4.85	11.68*	7.24	11.35*	2.22
	2	27.16*	18.31*	11.32*	3.26	28.82*	3.01	4.23	1.78
	3	55.02*	5.05	6.96*	0.88	19.40*	2.23	6.12	1.33
	4	16.21*	10.40*	7.74*	0.07	36.58*	5.16	23.64*	1.88
	5	48.08*	5.05	2.71	0.32	56.03*	1.09	8.23*	2.14
$m = 80$	1	77.54*	2.36	5.98	0.63	44.08*	3.37	10.86*	11.52*
	2	0.30	2.16	1.57	17.44*	95.03*	2.71	14.62*	10.78*
	3	50.91*	2.27	0.09	3.83	70.98*	4.39	26.91*	8.27
	4	36.58*	5.06	0.00	7.62	21.58*	11.06*	13.20*	13.57*
	5	20.99*	4.40	0.08	4.62	19.36*	12.42*	5.47	2.98

* $p < .01$.

inforced. On the other hand, repetitions were not selectively reinforced when the subjects tended to alternate. Switching always did better than repetitions.

An alternative approach to the frequency-dependent schedule is to identify the simplest response pattern that satisfies the constraints of the schedule and does better than any other competing pattern. In other words, borrowing an analogy from evolutionary biology, what is the *stable strategy* of the schedule? From the preceding analysis, it is clear that switching is that strategy because, when repeated continuously, switching ensures equal probabilities of reinforcement for each response on each trial. In variability schedules, however, switching is not stable: If alternations are rare, they will be differentially reinforced, but if they are dominant, they will be selectively extinguished. In fact, variability schedules have no simple stable strategy except random behavior (although counting in binary and cycling through all possible sequences is the most effective strategy, it is unlikely that pigeons and rats can learn it).

The question being asked in this study is how far in terms of pattern length one needs to go in order to generate random-like behavior. A one-response frequency-dependent schedule yielded only a first-order approximation to random performance (i.e., an equal

frequency of right and left pecks). It remains an open question whether the present schedule can produce more variability under slightly different conditions (e.g., longer intertrial intervals; Neuringer, 1991). Due to short-term memory constraints, for example, longer ITIs may prevent the development of sequential dependencies by attenuating the effects of the gradient of reinforcement for shift-and-stay responses evidenced in Figure 6. On the other hand, in order to generate variable or unpredictable responding without increasing the ITI, a frequency-dependent schedule will have to go beyond individual responses, at least to the level of pairs of key pecks. This possibility was explored in the next experiment.

EXPERIMENT 3

An analogy may be helpful in introducing the theoretical rationale guiding the third experiment. It is well known that some mathematical functions [e.g., $f(x) = e^x$] can be approximated by a power series expansion. Fast convergence of the series means that a good approximation can be obtained with only the first few terms of the expansion. Likewise, variable responding can be obtained with a frequency-dependent schedule acting on patterns made up of a small number of responses. The frequency-dependent schedule used in the

previous experiment was based on one-response patterns and generated a crude approximation to variable responding; in the present experiment, the frequency-dependent schedule was extended to two-response patterns.

In the one-response frequency-dependent schedule of Experiment 2, strict alternations constituted a stable strategy. Equivalently, in the two-response frequency-dependent schedule used in Experiment 3, double alternations (i.e., RRLRLRL...) constituted a stable strategy. If repeated continuously, double alternations yield an identical frequency of left and right pecks as well as of the four pairs LL, LR, RL, and RR (as before, pairs and longer response patterns were counted using a moving window). Due to the balancing nature of the schedule, this identical frequency of occurrence ensures an identical frequency of reinforcement of each individual response and pair. Whether pigeons could learn the double alternation pattern in the absence of external cues and with a probabilistic reinforcement rule was not known. Presuming they could not, it was predicted that they would adapt to the schedule through random responding (i.e., by varying their response patterns).

Finally, the problems addressed in Experiment 2 concerning sample size were also pertinent to this second-order schedule. Therefore, across the phases of the experiment, the sample size was manipulated.

METHOD

Subjects and Apparatus

Two pigeons from the previous experiment (157 and 156) and 2 experimentally naive pigeons (869 and 38) were used. Body weight and housing conditions were the same as in Experiment 2, and the same test chamber was used.

Procedure

The 2 naive birds learned to peck the keys under the autoshaping schedule described in Experiment 1. The experienced birds were exposed to one session of this procedure. Following this training, the transition phase was implemented for four sessions (see the procedure of Experiment 2). The number of one-response trials increased across sessions from 100 to 400, and the probability of reinforcement

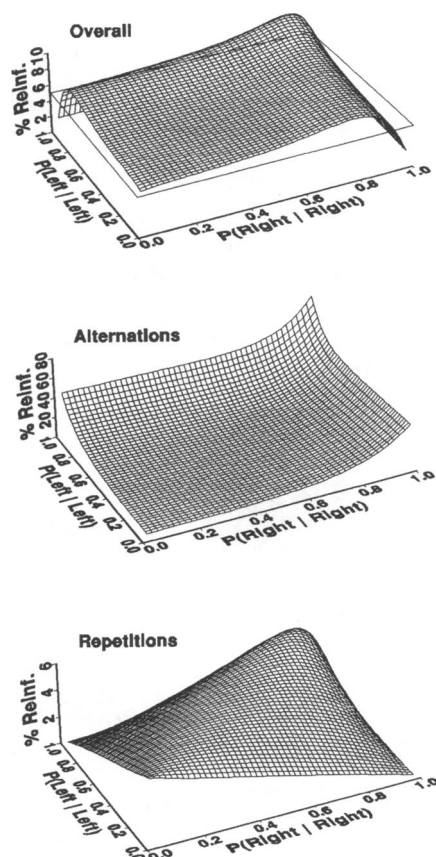


Fig. 6. Predicted probabilities of reinforcement when the response mechanism is a first-order Markov chain, the reinforcement rule is given by Equation 1, and a sample size of 10 is used. $P(R|R)$ and $P(L|L)$ are the two transition probabilities that characterize the chain. Notice that the vertical (z) axes have different scales. Top: Overall probability of reinforcement per trial, $P(\text{Ref})$. The plane $z = 5\%$ corresponds to the scheduled value, k , in Experiment 2. Middle: Probability of reinforcement given an alternation [i.e., $P(\text{Ref}|RL \text{ or } LR)$]. All values are above $z = 5\%$. Bottom: Probability of reinforcement given a repetition [i.e., $P(\text{Ref}|RR \text{ or } LL)$].

ment was gradually reduced to .08. Following the transition phase, the experiment proper began. All procedural details concerning the trial definition were exactly the same as in Experiment 2.

The reinforcement schedule. The reinforcement schedule worked as follows: From the last m pairs of responses, the probability of pecking the right key *given the last peck* was estimated. Thus, if the last peck was on the left key, then $P(R|L)$ equaled the number of times the pair LR occurred divided by the

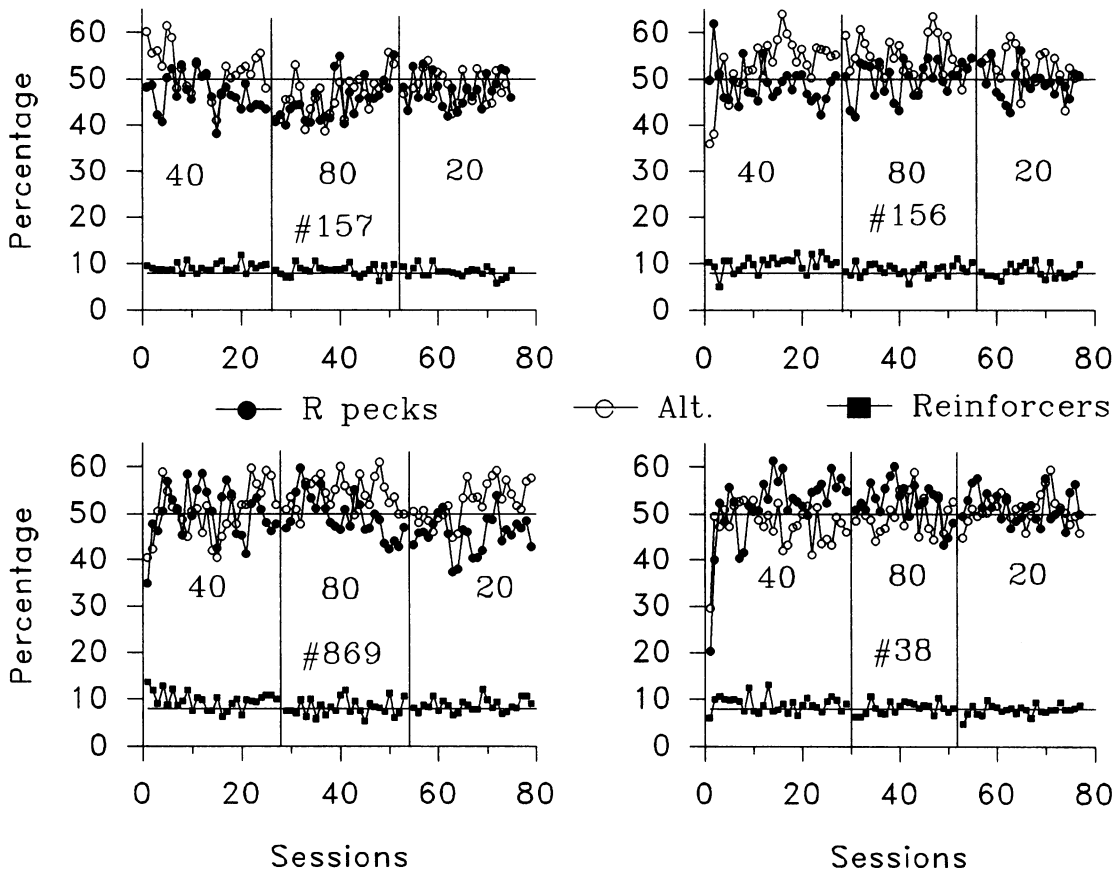


Fig. 7. Percentages of pecks on the right key, alternations, and reinforced pecks obtained in each session. The numbers in each condition show the sample size from which the probability of pecking the right key was estimated.

number of times the left response occurred: $P(R|L) = P(LR)/[P(LR) + P(LL)]$, because $P(LR) + P(LL) = P(L)$. Note that the pairs were counted with a moving window; in the string RRLR, for example, the pairs RR, RL, and LR would each be counted once. The payoff probabilities for the next trial were determined by Equation 1 using $P(R|L)$ as the s value. A similar procedure was followed when a right peck occurred: From the frequencies of the pairs RR and RL, the probability $P(R|R) = P(RR)/[P(RR) + P(RL)]$ was estimated and used as the s value in Equation 1. In the present experiment, s estimates a conditional probability and not, as in Experiment 2, an unconditional probability.

To clarify what the schedule attempts to accomplish, suppose that a bird has a strong tendency to switch keys, as in the preceding experiment. For such a bird, after a left peck

the estimate $P(R|L)$ will be close to one. In this case, the payoff probability for a left peck (i.e., a stay response) is substantially greater than the payoff probability for a right peck (i.e., a switch response). On the other hand, after a peck on the right key, the estimate $P(R|R)$ will be close to zero. In this case, $P(\text{Ref}|R)$ is much greater than $P(\text{Ref}|L)$. In short, when alternations are dominant, a repetition of the last response is differentially reinforced. Likewise, when repetitions are dominant, a switch response has a higher payoff probability. The only stable equilibrium occurs when shifts and stays are equally likely and, consequently, equally reinforced.

During the experiment proper, 400 one-response trials per session were conducted. In contrast with Experiment 2, no extinction trials were used to compute the initial response probabilities. Instead, at the beginning of each

Table 3

Chi-square values for a first- and second-order Markov chain model in the last five sessions of each condition. For a significance level of .01, critical values are $\chi^2(1) = 6.63$ and $\chi^2(2) = 9.21$, for tests of a first- and a second-order model, respectively. m = sample size.

Condi- tion	Session	Bird 157		Bird 156		Bird 869		Bird 38	
		1st	2nd	1st	2nd	1st	2nd	1st	2nd
$m = 40$	1	1.23	2.78	8.50*	2.68	7.27*	11.14*	4.17	2.21
	2	0.67	5.87	10.55*	4.68	3.64	5.94	3.27	13.90*
	3	3.92	0.62	7.06*	0.05	13.98*	14.08*	0.00	36.40*
	4	6.24	4.46	4.00	0.82	11.71*	24.49*	0.11	24.90*
	5	0.18	0.18	5.29	5.50	1.08	13.58*	2.34	7.90
$m = 80$	1	0.03	1.12	2.92	0.97	7.32*	1.93	3.86	9.93*
	2	0.63	5.13	1.02	5.99	2.90	10.28*	2.99	14.17*
	3	0.17	1.96	0.63	1.58	3.71	3.69	1.34	6.85
	4	5.47	7.89	0.92	4.48	0.28	4.38	0.53	5.29
	5	3.01	0.62	4.02	2.20	0.03	5.61	1.54	7.79
$m = 20$	1	4.24	5.05	0.17	0.29	3.58	5.75	0.38	6.53
	2	0.37	0.16	6.65*	1.94	0.71	4.31	0.12	22.73*
	3	0.92	1.43	1.56	0.12	0.34	2.77	0.39	14.27*
	4	0.08	3.22	0.06	1.11	7.40*	2.38	0.03	5.05
	5	2.45	2.23	0.17	0.29	13.49*	11.72*	2.87	24.82*

* $p < .01$.

session, the s value was computed from the previous trials including, when necessary, the last trials of the previous session. All other details remained identical to Experiment 2.

In the three phases of the experiment, the sample size equaled 40, 80, and 20, and the expected probability of reinforcement per trial, k , equaled .08, .08, and .06, respectively. When the sample size was 20, the probability of reinforcement was reduced to .06 because computer simulations had shown that when small samples are used, the obtained probability of reinforcement is slightly greater than the scheduled value.

RESULTS AND DISCUSSION

Data analysis was based on the string of 400 individual pecks emitted by each subject in each session. As in Experiments 1 and 2, pairs, triplets, and, in general, n -tuplets of responses were counted using a moving window.

Figure 7 displays the percentage of pecks on the right key, the percentage of pairs involving an alternation (i.e., RL + LR), and the percentage of reinforced pecks for each bird in each session. As in Experiment 2, irrespective of the sample size m , the percentage of pecks on the right key was always close to 50%, with most values lying between 40% and 60%. Alternations were also generally close to 50% in all conditions. For Birds 157, 156, and 38,

the percentages of pecks on the right key and alternations showed less between-sessions variance when the sample was 20 than before. These results indicate that the schedule attained the goal of balancing not only right and left pecks but alternations and repetitions as well.

Lag analysis and Markov tests were used to determine if any higher order dependencies were developed. Figure 8 shows the lag data for each bird in the last five sessions of each condition, and Table 3 presents, for the same sessions, the chi-square values that tested the Markovian model. Because experienced and naive birds performed differently, the analysis for each of the two groups is described separately.

The lag data in Figure 8 reveal that the two experienced birds, 157 and 156, gradually eliminated the alternation pattern clearly seen in some sessions when the sample size equaled 40 and 80. At the end, when the sample was 20, no strong or reliable sequential dependencies were obvious. In the Markov analysis in Table 3, no chi-square test was significant for Bird 157; for Bird 156, a first-order dependency occurred in three of the last five sessions when the sample size was 40 and once when the sample was 20. These results suggest that, with training, the 2 experienced birds approached random-like performance. To fur-

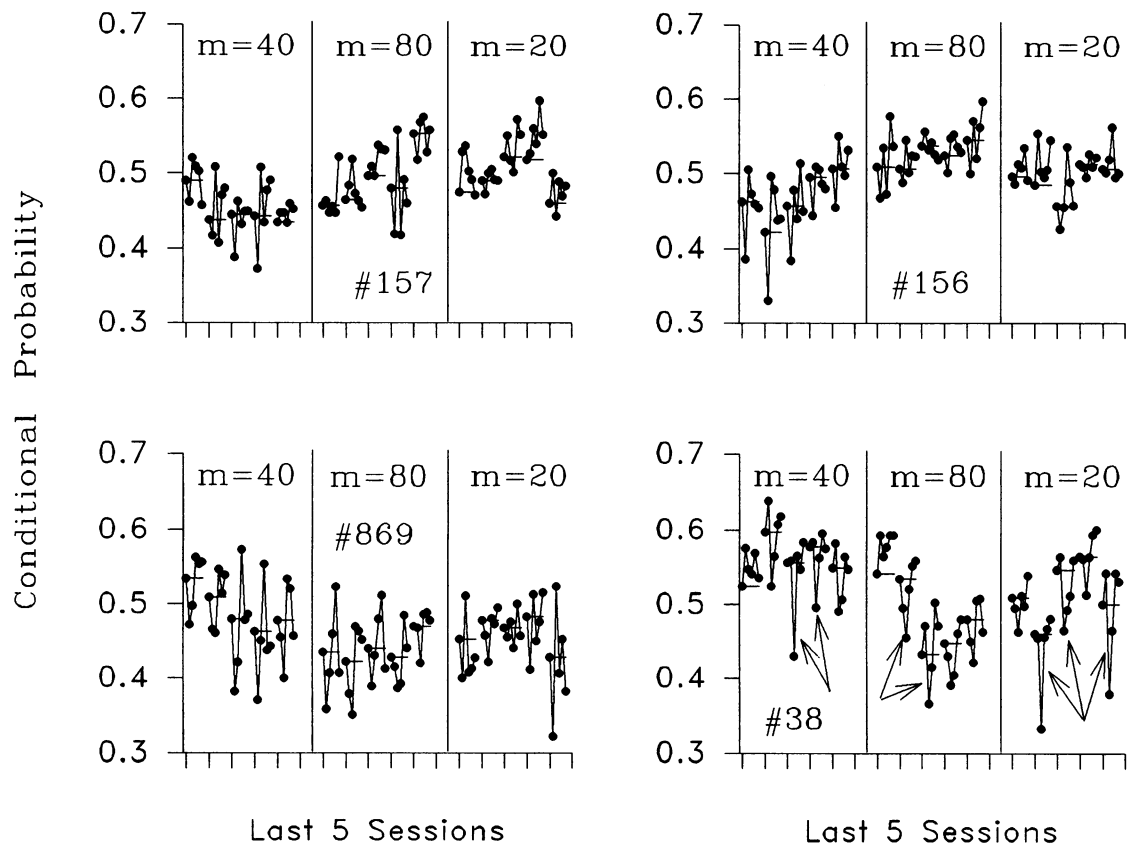


Fig. 8. Conditional probability profiles for the last five sessions of each condition. Values of m refer to the sample size used during each condition. The arrows for Bird 38 show profiles that correspond to a high frequency of the double alternation pattern RRL... Other details as in Figure 2.

ther substantiate this conclusion, Figure 9 shows within-session analyses of the frequency of right pecks and runs of right pecks for the last five sessions of Bird 157 when $m = 20$. To assess the within-session stability of the probability of pecking right, each session was divided into 10 blocks of 40 responses each. The frequency of right pecks in each block is shown in Figure 9, with the solid line connecting the mean of the five sessions. The frequency of right pecks was relatively constant and close to 20, as expected from a Bernoulli process with parameter $P(R) = .5$.

A run of length x corresponds to a string of x consecutive right pecks preceded and followed by a left peck. Thus, the string LRRRL would count as a run of Length 3. Figure 9 shows that the frequency of each run was also very close to the value predicted by a Bernoulli process with parameter $P(R) = .5$.

The 2 naive birds, 38 and 869, displayed less response variability. Thus, in several sessions of each condition, the lag data for Bird 38 showed a profile with Lag 2 much lower than Lag 0, and Lag 1 close to Lag 0 (see arrows in Figure 8 for examples of this profile). This profile is expected if the stable strategy, the RRL pattern, is performed with added noise. The chi-square tests of Markov models confirmed the development of second-order patterns: Eight of 15 tests with Bird 38 suggested a second-order dependency, but no test suggested a first-order or a third-order dependency. Table 3 shows the chi-square tests for the first- and second-order models.

When a moving window of Size 4 is used to count sequences, the repeated presentation of the double alternation pattern (i.e., RRLRLRL...) yields an increased frequency of the four sequences RRL, RLLR,

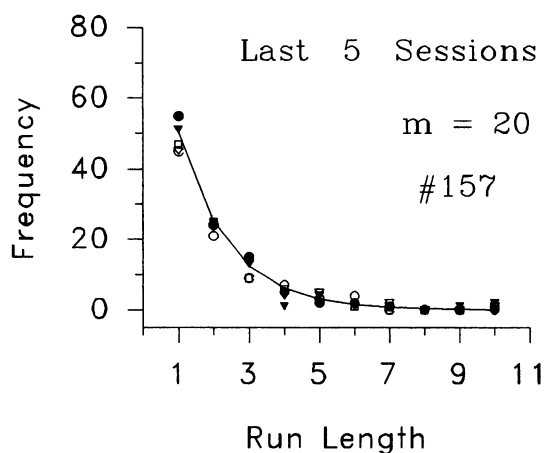
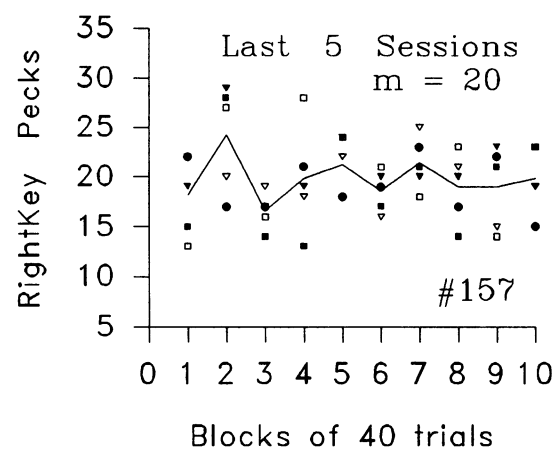


Fig. 9. Top: Frequency of right pecks in each block of 40 responses for the last five sessions of Bird 157 when $m = 20$. The solid line connects the mean of the five sessions. Bottom: Frequency of runs of right pecks for the same bird during the same sessions. The solid line connects the predicted values according to a Bernoulli process with $P(R) = .5$.

LLRR, and LRRL. This result was frequently observed; Figure 10 shows a typical session at $m = 40$ for Bird 38 in which the stable pattern clearly dominated.

For Bird 869, the lag profile corresponding to the pattern RLLRLL... was displayed frequently during the first two phases (Figure 8). In the last phase, when the sample size was 20, a pattern of simple alternations was sometimes dominant. The chi-square tests (Table 3) suggested a second-order model in 6 of 15 sessions; in two sessions, only a first-order dependency was observed. Again, no test suggested a third-order model. Hence, for Bird

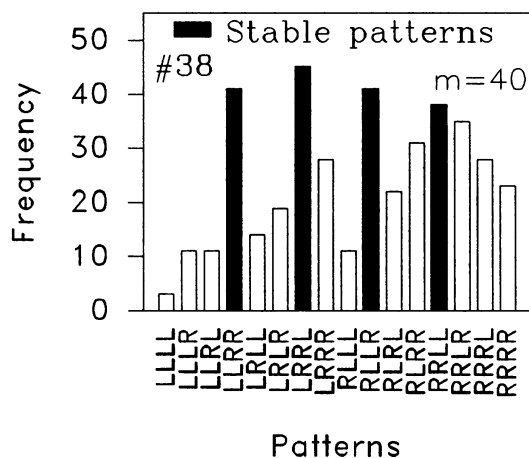


Fig. 10. Frequency of each sequence of four responses produced by Bird 38 during one session when $m = 40$. The filled bars correspond to the four-response sequences associated with the stable double alternation pattern (i.e., RLLRLL...).

869, either the preceding response or the last two responses affected the probability of the next peck.

In summary, there was a clear indication that naive birds developed second-order patterns. Consequently, their performance differed from that observed under variability schedules. On the other hand, with training, the experienced birds approached random, or Bernoulli-like, behavior. Their performance matched that observed under variability schedules, which confirms the hypothesis that variable responding can result from a frequency-dependent selection process that targets only two-response patterns.

Due to the small number of experienced and naive subjects used in this experiment, the following conclusions must be taken cautiously. Figure 7 shows that naive and experienced birds obtained similar numbers of reinforcers per session. Therefore, it was not the case that one response pattern (e.g., double alternation) was more or less effective than random responding. On the other hand, it is possible that previous exposure to a schedule that selected strict alternations (Experiment 2) prevented the experienced birds from learning the stable pattern during the current experiment. This hypothesis is reminiscent of Schwartz's (1982b) findings that previous exposure to contingent reinforcement may impair the ability to discover new rules in subsequent situations.

In order to engender variable performance in naive birds, a frequency-dependent schedule would probably need to target three-response patterns, but not longer ones. The stable strategy in the three-response case, RRRLRLLL, is not likely to be learned by pigeons. (To see that RRRLRLLL is the stable strategy, note that, when continuously repeated, this pattern yields an identical frequency of occurrence of each triplet. Then, under the frequency-dependent schedule, each triplet, pair, and single response would be equally reinforced.) Also, the second-order patterns developed by the naive birds were not as strong as the switching pattern developed by all birds in Experiment 2. Finally, in some sessions the naive birds did display close-to-random behavior, as revealed by the lag data and the Markov tests. Experiment 3 may have approached the limits of pigeons' capabilities to learn extended response patterns.

GENERAL DISCUSSION

In Experiment 1, when reinforcers were delivered after variable sequences of four pecks, response variability increased: The proportions of right pecks and alternations approached .5, few sequential dependencies were shown, and the response mechanism was formally equivalent to a Bernoulli process. When the same probability of reinforcement was scheduled irrespective of sequence variation, response variability was greatly reduced.

The reinforcement rule in variability schedules is time (or trials) related, in that a sequence is reinforced provided it did not occur during the last N trials. This rule, it was suggested, works through the selective reinforcement of the momentarily least probable patterns. Probability, not time, number of trials, or temporal order, became the central variable. Thus, the general hypothesis under study ascribed the operant conditioning of response variability to a process of frequency-dependent selection.

The empirical test of this hypothesis is necessarily indirect because an additional specification is needed, namely the target of selection. A rejection of the hypothesis in one experiment may simply indicate an erroneous identification of the appropriate target. In Experiment 2, the simplest possibility, a one-re-

sponse frequency-dependent schedule, was studied. It was assumed that the birds' response mechanism was Bernoulli-like, and the schedule simply corrected any deviations away from the indifference point $P(R) = .5$. It was found that the schedule continuously favored shift over stay responses, and the birds behaved accordingly (i.e., they alternated frequently). In other words, the response mechanism developed into a first-order Markov chain with high switching probabilities, $P(R|L)$ and $P(L|R)$. Hence, only a first-order approximation to variable performance, the identical frequency of right and left pecks, was obtained.

Experiment 3 assumed a first-order Markov chain as the response mechanism and extended the frequency-dependent selection process to pairs of responses. The schedule attempted to correct any deviations away from the indifference points $P(R|R) = .5$ and $P(R|L) = .5$. If this correction procedure is successful [i.e., if $P(R|R) = P(R|L) = .5$] and, in addition, no higher-order patterns are developed, then the Markov chain reduces to the Bernoulli process and maximal response variability is obtained.

The results showed that naive birds developed second-order response units. One of them learned the double alternation stable pattern, and, consequently, its behavior showed a second-order approximation to random performance (equal preference for left and right pecks and for each response pair). With training, the experienced birds came to display high degrees of behavioral variability, comparable in all respects to that observed under variability schedules.

These results suggest, but do not prove, that variable responding is obtained when the following conditions (some necessary, others sufficient) are met. Some form of frequency-dependent or balancing selection is most likely a necessary condition; that is, the reinforcement rule must somehow select the response that is momentarily less likely to occur. Two other lines of evidence support this hypothesis. First, most (if not all) animal studies that successfully conditioned response variability have used some form of frequency-dependent selection (Blough, 1966; Bryant & Church, 1974; Machado, 1989; Morris, 1987; Neuringer, 1991; Page & Neuringer, 1985; Pryor et al., 1969; Schoenfeld et al., 1966; Van Hest et al., 1989). Second, other schedules of reinforcement that

typically engender partial preferences (a necessary condition to obtain response variability) also implement some "mild" form of frequency-dependent selection. The most well-known example is the concurrent variable-interval variable-interval schedule in which the probability of reinforcement for one alternative increases with time since the last response to that alternative (e.g., Hinson & Staddon, 1983a, 1983b). An experiment that showed the operant conditioning of response variability without frequency-dependent selection would certainly disconfirm the claim that frequency-dependent selection is a necessary condition, but, to my knowledge, that experiment has never been done.

Although possibly necessary, frequency-dependent selection per se is not a sufficient condition because the unit or target of selection also plays a crucial role. If the unit is very simple (e.g., individual key pecks in pigeons), additional constraints (e.g., longer ITIs; Fetterman & Stubbs, 1982; Neuringer, 1991) may also be needed to prevent the development of higher order stable patterns. Alternatively, a frequency-dependent schedule targeting longer response units may be a sufficient condition because the corresponding higher order stable strategies are too complex to be learned. With pigeons and their key-pecking responses, it is possible that these units are not longer than three key pecks (all birds alternated reliably in Experiment 2 but only one developed double alternations in Experiment 3).

The ultimate fate of the frequency-dependent selection hypothesis remains a matter of future theoretical and empirical investigations. In order to pave the way for such investigations, the remainder of this section considers the behavioral mechanisms assumed by this hypothesis and compares them with alternative viewpoints.

A general problem posed by the very conceptualization of variability as an operant was raised by Schwartz (1982a, p. 78) when he asked "What objective property of responses would unite them into a class?" If one assumes that reinforcement only *strengthens* behavior, then how could variability be directly strengthened if it does not correspond to any objective property of responses? Later, Page and Neuringer (1985) advanced a two-component hypothesis that emphasized the *shaping* effects of rewards: (a) Pigeons discriminate between

contingencies in which reinforcement follows repetitions and contingencies in which reinforcement requires nonrepetitions; (b) in the latter condition, an inborn variability generator would be turned on and tuned to the current variability requirement (Neuringer, 1986, p. 74; Page & Neuringer, 1985, pp. 449–450). This hypothesis contains important omissions and a major conceptual difficulty of its own. Although the nominal contingency in variability schedules specifies response sequences with fixed length, I have argued that these schedules target simultaneously patterns of various lengths. Hence, what exactly is being conditioned and randomly generated: individual responses, such as left and right pecks, or extended patterns of responses? If the discrimination process mentioned in (a) is based on sequences, how long are these sequences? More seriously, however, point (b) assumes that pigeons are *directly* sensitive to the variability criterion N (assumed to fine tune the random generator). If this were the case, then it is not clear why pigeons would ever turn on a random generator instead of responding systematically, cycling through $N + 1$ different sequences when the reinforcement criterion is set to N , and collect all the available reinforcers. The reply that memory limitations preclude systematic response variation would be circular, because the initial assumption of direct sensitivity to N itself presupposes the absence of those limitations.

The frequency-dependent selection hypothesis assumes that reinforcement strengthens *and* shapes responses. The shaping effect is revealed in the formation of new response units from the building blocks of right and left responses (e.g., alternations and double alternations in Experiments 2 and 3, respectively). The strengthening effect is revealed by the relative frequency of each response unit. To be sure, the hypothesis also assumes a random generator that, metaphorically speaking, scans the available alternatives and selects each one with a probability proportional to its current strength. However, the response alternatives themselves can be changed by the reinforcement rule. Variable performance is predicted only when (a) no higher order stable response unit is, or can be, effectively learned, and (b) when all the current response units are equally strong due to the balancing nature of the frequency-dependent schedule.

Another hypothesis (e.g., Neuringer, 1991) emphasizes memory disruption as the mechanism engendering response variability. Stated simply, the hypothesis claims that in order to behave unpredictably in the present, we must forget what we did in the past. However, unless we assume that variability is the default behavioral state when memory is disrupted, it is not clear why we should vary our responses when the past is forgotten: Why should we not repeat them? A strongly biased coin does not remember its past, but this does not prevent one side of the coin from occurring all the time. The memory hypothesis must be supplemented with a *decision rule* with clear implications for action when memory is in fact disrupted.

The frequency-dependent selection hypothesis does not deny a role for memory; rather, it redefines this role. Memory is seen as a constraint on pattern formation. As the frequency-dependent selection process is extended to larger response units, the corresponding stable patterns become more difficult to learn and the likelihood of random responding is thereby increased. Long interresponse intervals (Neuringer, 1991) may achieve the same goal of precluding the acquisition of higher order patterns. The generation of performance characterized by the identical frequency of each alternative, however, is mainly due to frequency-dependent or balancing selection.

The terms *randomness* and *variability* are sometimes treated as synonyms. However, variable behavior can be achieved by nonrandom, deterministic rules, and randomness can entail nonvariable, stereotypical responding, as when the parameters of a stochastic process take extreme values. Despite the difficulties associated with the concept of randomness (e.g., Gigerenzer et al., 1989; Lopes, 1982) these two possibilities imply that variability and randomness describe properties of different objects. Randomness is a property of a *process* or *mechanism*, like the response-generating mechanism. Variability, on the other hand, is a property of the *outcome* of the process or mechanism. Numbers in a table or a pigeon's key pecks are not, properly speaking, random, but more or less variable or predictable; we can determine the likelihood that they were generated by a particular random process (this is the hallmark of statistical inference) or, con-

versely, we can try to find out the random process that may have engendered them.

This distinction between randomness and variability may throw new light onto a current argument concerning reinforcement and variability. The study of this topic was boosted by the "negative" results obtained by Schwartz (1980, 1982a), which led him to conclude that reinforcement could generate only stereotyped behavior, and the experiments carried out by Page and Neuringer (1985) and others reaching opposite conclusions. Since then, it has been concluded (e.g., Catania, 1987) that the task Schwartz used contained an artifact that prevented pigeons from behaving variably. In that task, pigeons were required to peck exactly four times on each response key in order to displace a light from the left upper corner to the lower right corner of a 5×5 square matrix of lights; a peck on the right key displaced the light one position down, and a peck on the left displaced it one position to the right. Once the current sequence had to differ from the previous one, pigeons performed poorly, in clear contrast with their performance in the absence of any variability requirement. Page and Neuringer (1985) argued that this result was due to the combinatorics of the task: Although eight pecks, distributed over two keys, can occur in 256 different ways, the constraint of four pecks per key reduces the possibilities to 70. However, the birds in Schwartz's experiments lost most rewards not by behaving randomly but by repeating the sequences that had previously been systematically reinforced. As Schwartz (1980) himself argued, the problem might have been due to the response integration that occurred during preliminary training; in a sense, the birds' repertoire had been severely restricted to the few (sometimes only one) sequences developed initially. What pigeons could not do, at least under the circumstances, was to vary their sequences *systematically*, that is, within the limits imposed by the matrix (see Schwartz, 1988). These difficulties are not so much related to reinforcement per se, as Schwartz concluded, as to memory (of the complex sequences leading to the goal and of the previous sequence that would allow the bird to avoid it on the next trial). It is possible that complicated sequences, whose performance depends on external cues, cannot be randomly generated. Can pigeons learn an equivalent task that also requires systematic variation but

uses simpler patterns? The boundary conditions under which reinforcement can engender systematic, as opposed to random-like, variation remain to be determined.

REFERENCES

- Bakeman, R., & Gottman, J. (1986). *Observing interaction: An introduction to sequential analysis*. New York: Cambridge University Press.
- Blough, D. S. (1966). The reinforcement of least-frequent interresponse times. *Journal of the Experimental Analysis of Behavior*, 9, 581-591.
- Boulanger, B., Ingebo, A., Lahak, M., Machado, A., & Richelle, M. (1987). Variabilité comportementale et conditionnement opérant chez l'animal [Behavioral variability and operant conditioning in animals]. *L'Année Psychologique*, 87, 417-434.
- Bryant, D., & Church, R. M. (1974). The determinants of random choice. *Animal Learning & Behavior*, 2, 245-248.
- Catania, A. C. (1987). Some Darwinian lessons for behavior analysis: A review of Bowler's *The Eclipse of Darwinism*. *Journal of the Experimental Analysis of Behavior*, 47, 249-257.
- Fetterman, J. G., & Stubbs, D. A. (1982). Matching, maximizing, and the behavioral unit: Concurrent reinforcement of response sequences. *Journal of the Experimental Analysis of Behavior*, 37, 97-114.
- Gigerenzer, G., Swijtink, Z., Porter, T., Daston, L., Beatty, J., & Kruger, L. (1989). *The empire of chance: How probability changed science and everyday life*. New York: Cambridge University Press.
- Gottman, J. M., & Roy, A. K. (1990). *Sequential analysis*. New York: Cambridge University Press.
- Hinson, J. M., & Staddon, J. E. R. (1983a). Hill-climbing by pigeons. *Journal of the Experimental Analysis of Behavior*, 39, 25-47.
- Hinson, J. M., & Staddon, J. E. R. (1983b). Matching, maximizing, and hill-climbing. *Journal of the Experimental Analysis of Behavior*, 40, 321-331.
- Lopes, L. L. (1982). Doing the impossible: A note on induction and the experience of randomness. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 8, 626-636.
- Machado, A. (1989). Operant conditioning of behavioral variability using a percentile reinforcement schedule. *Journal of the Experimental Analysis of Behavior*, 52, 155-166.
- Morris, C. J. (1987). The operant conditioning of response variability: Free-operant versus discrete-response procedures. *Journal of the Experimental Analysis of Behavior*, 47, 273-277.
- Morris, C. J. (1989). The effects of lag value on the operant control of response variability under free-operant and discrete-response procedures. *The Psychological Record*, 39, 263-270.
- Neuringer, A. (1986). Can people behave "randomly"? The role of feedback. *Journal of Experimental Psychology: General*, 115, 62-75.
- Neuringer, A. (1991). Operant variability and repetition as functions of interresponse time. *Journal of Experimental Psychology: Animal Behavior Processes*, 17, 3-12.
- Page, S., & Neuringer, A. (1985). Variability is an operant. *Journal of Experimental Psychology: Animal Behavior Processes*, 11, 429-452.
- Platt, J. R. (1973). Percentile reinforcement: Paradigms for experimental analysis of response shaping. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in theory and research* (Vol. 7, pp. 271-296). New York: Academic Press.
- Pryor, K. W., Haag, R., & O'Reilly, J. (1969). The creative porpoise: Training for novel behavior. *Journal of the Experimental Analysis of Behavior*, 12, 653-661.
- Schoenfeld, W. N., Harris, A. H., & Farmer, J. (1966). Conditioning response variability. *Psychological Reports*, 19, 551-557.
- Schwartz, B. (1980). Development of complex, stereotyped behavior in pigeons. *Journal of the Experimental Analysis of Behavior*, 33, 153-166.
- Schwartz, B. (1982a). Failure to produce response variability with reinforcement. *Journal of the Experimental Analysis of Behavior*, 37, 171-181.
- Schwartz, B. (1982b). Reinforcement-induced behavioral stereotypy: How not to teach people to discover rules. *Journal of Experimental Psychology: General*, 11, 23-59.
- Schwartz, B. (1988). The experimental synthesis of behavior: Reinforcement, behavioral stereotypy, and problem solving. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in theory and research* (Vol. 22, pp. 93-138). New York: Academic Press.
- Shimp, C. P. (1967). Reinforcement of least-frequent sequences of choices. *Journal of the Experimental Analysis of Behavior*, 10, 57-65.
- Van Hest, A., Van Haaren, F., & Van De Poll, N. E. (1989). Operant conditioning of response variability in male and female Wistar rats. *Physiology and Behavior*, 45, 551-555.

Received October 21, 1991

Final acceptance April 10, 1992